MAKERERE UNIVERSITY

COLLEGE OF COMPUTING AND INFORMATICS TECHNOLOGY

AUTOMATION OF CUSTOMER SUPPORT IN THE TELECOM INDUSTRY USING MACHINE LEARNING

By

CS22-24

DEPARTMENT OF COMPUTER SCIENCE

SCHOOL OF COMPUTING AND INFORMATICS TECHNOLOGY

A Project Proposal Submitted to the School of Computing and Informatics Technology for the Study Leading to a Project Report in Partial Fulfillment of the Requirements for the Award of the Degree of Bachelor of Science in Computer Science of Makerere University

Supervisor
Mr. Kizito Jonathan
School of Computing and Informatics Technology
jkizito@cis.mak.ac.ug, +256-414-540-628

October, 2022

TABLE OF CONTENTS

DECLARATION **i**

APPROVAL **ii**

ACKNOWLEDGEMENT **iii**

LIST OF FIGURES **iv**

LIST OF TABLES **v**

GLOSSARY **vi**

ABSTRACT **vii**

1 INTRODUCTION **1**
1.1 BACKGROUND 1
1.2 PROBLEM STATEMENT 1
1.3 RESEARCH OBJECTIVES AND QUESTIONS 2
1.3.1 MAIN OBJECTIVE 2
1.3.2 SPECIFIC OBJECTIVES 2
1.3.3 RESEARCH QUESTIONS 2
1.4 RESEARCH SCOPE 2
1.4.1 GEOGRAPHICAL SCOPE 2
1.4.2 TECHNICAL SCOPE 2
1.5 TIME SCOPE 3
1.6 RESEARCH SIGNIFICANCE 3

2 LITERATURE REVIEW **4**
2.1 DEFINITION OF THE CUSTOMER TROUBLE TICKET 4
2.2 DEFINITION OF THE TROUBLE TICKET SYSTEM 4
2.3 THE LIFECYCLE OF A CUSTOMER TROUBLE TICKET 5
2.3.1 TRANSITION STATES OF A TROUBLE TICKET 5
2.3.2 TRANSITION PROCESSES OF A TROUBLE TICKET 5
2.4 CUSTOMER SUPPORT KEY PERFORMANCE INDICATORS 7
2.5 CUSTOMER SUPPORT IN THE TELECOM INDUSTRY 9
2.5.1 KEY CHANNELS OF CUSTOMER SUPPORT IN-USE 9
2.5.2 CASE-STUDY: TWITTER AS A TROUBLE TICKET SYSTEM CHANNEL . . 10
2.5.3 MERITS OF TWITTER AS A TROUBLE TICKET SYSTEM CHANNEL 11
2.5.4 DEMERITS OF TWITTER AS A TROUBLE TICKET SYSTEM CHANNEL . . 11
2.6 TICKET CLASSIFICATION THROUGH SENTIMENT ANALYSIS 12
2.6.1 LITERATURE REVIEW ON SENTIMENT ANALYSIS 12
2.6.2 MODE OF APPLICATION 12

3 RESEARCH METHODOLOGY **14**
3.1 INTRODUCTION 14
3.2 RESEARCH DESIGN 14
3.3 AREA OF RESEARCH STUDY 14

3.4 STUDY POPULATION AND SAMPLING TECHNIQUES 14
3.5 METHODS OF DATA COLLECTION 14
3.5.1 ONLINE SURVEY 14
3.5.2 WEB SCRAPING 15
3.6 INSTRUMENTS OF DATA COLLECTION 15
3.6.1 GOOGLE FORMS 15
3.6.2 TWITTER API 15
3.7 RELIABILITY AND VALIDITY OF RESEARCH INSTRUMENTS 16
3.8 METHODS OF DATA ANALYSIS 16
3.8.1 DATA WRANGLING 16
3.8.2 EXPLORATORY DATA ANAYLYSIS 17
3.8.3 EXPLANATORY DATA ANALYSIS AND VISUALIZATION 17
3.9 ETHICAL CONSIDERATIONS 17

4 RESEARCH RESULTS **18**
4.1 INTRODUCTION 18
4.2 DEMOGRAPHIC DISTRIBUTION OF RESPONDENTS 18
4.3 PREFERENCES OF RESPONDENTS 19
4.4 ANALYSIS OF KPIs ON TICKETS THROUGH TWITTER 23
4.5 ANALYSIS OF CHOSEN NLP LIBRARIES 25
4.5.1 EFFICIENCY OF MODELS IN MINING ASPECTS 25
4.5.2 ANALYSIS OF TICKET- SUBJECTIVITY AND POLARITY 27
4.5.3 ACCURACY OF NLP LIBRARIES 27

5 DISCUSSION OF RESULTS **29**
5.1 EXPLANATION OF RESEARCH RESULTS 29
5.2 CHALLENGES AND LIMITATIONS 30
5.3 HYPOTHESES BASED ON RESEARCH RESULTS 30

6 CONCLUSION **31**

7 RECOMMENDATIONS **32**

8 REFERENCES **33**

DECLARATION

We, the undersigned authors, hereby declare that the this project report has been solely composed by us and that this work has not been submitted for any other degree or professional qualification. We confirm that the work hereby submitted is our own, and where publications were citated, due references have been provided. This study was jointly conceived by all the undersigned authors.

No.	Names	Registration Number	Student number	Signature
1.	BUBUKA SHARIF	19/U/16355/EVE	1900716355	
2.	JEMBA TONY	18/U/25544/PS	1800725544	
3.	ENGENA JEROME BRIAN	19/U/16418/EVE	1900716418	
4.	NAMUTEBI MARY BRENDA	19/U/17451/EVE	1900717451	

APPROVAL

I, hereby recommend that this project report entitled "AUTOMATION OF CUSTOMER SUPPORT IN THE TELECOM INDUSTRY USING MACHINE LEARNING", prepared and conceived by Bubuka Sharif, Jemba Tony, Engena Jerome Brian and Namutebi Mary Brenda under my supervision be accepted in partial fulfillment of the authors' requirements for the award of their degrees of Bachelor of Science in Computer Science at Makerere University.

Signed: ....

Date: ....11/10/2022..........................

MR. KIZITO JONATHAN

MAKERERE UNIVERSITY

ACKNOWLEDGEMENT

It has been a long and strenuous journey, but we have finally reached the finish line with our final year project. We owe an immense amount of gratitude to all those who contributed in one way or another; without you, this would not be possible. The completion of such a large-scale endeavor requires support from many people, so thank you for your contributions!

In a special way however, we are very grateful to our parents and teachers across all levels of education we have gone through, who together have toiled so that we merit and be able to present such an incredible piece of work. Thank you to our parents who have paid our school dues for all these years, and endured all forms of economic turmoil,to have us reach this far as far as education is concerned. Likewise, we are grateful to our teachers, especially our lecturers at Makerere university, who have always been helpful beyond limits. Thank you very much!

Finally, we want to extend our thanks towards God (The Great Almighty) because he always had his blessing on us when nobody else did.

LIST OF FIGURES

1	States of a customer trouble ticket in a trouble ticket system	5
2	A customer support ticket to Airtel on Twitter	10
3	Exploratory analysis of a few columns of the survey responses datset	17
4	Gender distribution of respondents	18
5	Primary telecom subscriptions of respondents	18
6	Social media channel accounts owned by respondents	19
7	Preffered channels for customer support in the telecom industry by respondents	19
8	Satisfaction with customer suupport on Twitter by the respondents	20
9	Challenges faced by customers during customer support sessions	20
10	Most important factor during a customer support session	21
11	A histogram showing acceptance of chatbots among respondents	21
12	Most essential features in a trouble ticket system	22
13	A bar graph of the necessity of a dedicated unichannel customer support platform	22
14	Average Reply Time (ART) of MTN and AIRTEL	23
15	Tickets that were best suited for a forum instead of live twitter support	23
16	Tickets suited for a self-help knowledge repository instead of live twitter support	24
17	A distribution of the different types of tickets that were analyzed	24
18	A distribution of the support levels that were required by the analyzed tweets	25
19	A correlation between an agents total ticket count and their average response time	25
20	The reduction in aspects as seen by an agent pre- and post-ABSA	26
21	The 4 most common aspects in the 1003 tickets	26
22	A correlation of ticket subjectivity and ticket polarity	27
23	Classification of tickets based on mood sentiment (TEXTBLOB)	28
24	Performance of TextBlob in classification of tickets	28

LIST OF TABLES

1 Summary of the categories of trouble ticket systems 4
2 Summary of the five levels of customer support. 6
3 Categories of employees in a eNPS index. 8
4 Sample results of PyABSA and TextBlob 13

GLOSSARY

anticipatory ticket This is a would-be support ticket that is predicted through anlyzing the lifecycle of a customer, and customer is proactively innitiated to them before ticket creation.. 4

customer attrition Also known as customer churn, customer turnover, or customer defection, this is the loss of clients or customers by a company.. 9

omnichannel trouble ticket system This is a multichannel based trouble ticket system — one that supports the creation of trouble tickets from various disparate external systems like emails, telephoney, social media platforms, websites, applications among others.. 5, 11

responsive ticket This is a ticket created post experiencing a service problem or incident by the customer, and customer support can not have predicted its creation.. 4

service-level agreement This is a formal documented commitmet between a service provider and a client, that clearly lays out the metrics by which the service is measured, as well as remedies or penalties agreed-on in case service levels are not achieved.. 7

unichannel trouble ticket system This is a single channel based trouble ticket system — one that supports the creation of trouble tickets only from a selected group of sources, usually internal, such as a dedicated websites, mobile app, telephone line, or email address.. 5

ABSTRACT

The purpose of this study was to establish the functional loopholes in the customer support rendered in the telecom industry in Uganda, with a major focus on Twitter as the support channel, in order to assess how best the processes therein can be automated using machine learning in order to improve the overall customer experience.

In this research, we considered a descriptive research design, and an endeavour was taken to use both qualitative and quantitative data to establish the state of customer support in the telecom industry in Uganda, as well as how best machine learning can be leveraged to improve it through automation. The target population consisted of 155 volunteers and the two biggest telecom service providers in Uganda by market share, that is — MTN and AIRTEL. Random sampling was used to select the online survey respondents, who were majorly students at Makerere University. 1003 customer support tweets sent to MTN and AIRTEL before 15th September 2022 were programmatically collected and analyzed to assess the the performance of these companies across various metrics in customer support, as well as to validate and justify our proposed method of automation.

An online survey and the Twitter API were employed during the collection of data. Microsoft Excel was then extensively used to clean, code and enrich this data. Python was the primary language and tool used in the analysis of the data. The Numpy and Pandas libraries were used in the exploratory analysis of the data and Matplotlib was used in the explanatory, or rather visualization phase of the data analysis. When correlations were required, the Pearson correlation coefficient was used.

After extensive research through existing literature, sentiment analysis using machine learning was proposed as a solution to automate some processes in the pipeline of a trouble ticket system. Two libraries, that is — PyABSA and TextBlob were used to analyze and justify this proposal. The research findings to a large degree proved the necessity for handling customer support in a platform or environment dedicated to that purpose, since such environments offer the best opportunities to automate processes in trouble ticket systems.

The research paper concludes with a few recommendations on how best this research can be used to not only support subsquent research in this area, but also guide stakeholders concerned with customer care, support and experiene across various domains. Specifically, further research is encouraged in the automation of processes involved in e-governance, since it has one of the biggest amounts of data and highest need for automation.

1 INTRODUCTION

1.1 BACKGROUND

In his book published in 2012, Robert W. Lucas described customer support as the assistance provided by a company to those people who buy or use its products or services [Lucas, 2012]. Customer care services could include but not be limited to practices such as offering reliable services, top-notch security, front-desk services, speed of customer service, customization of service delivery, honesty among others. In most cases, customer support is an important factor in the ability of a company to retain customers. To a young and rapidly growing entreprise, good customer support could even be used an as economic moat — a competitve advantage to preserve or grow its market share as well as increase its profit margins.

Since the advent of online commerce (Ecommerce), it has been increasingly expected of any business by its customers to be able to offer fast and efficient customer support to them, wherever they are and whenever they need it. Increasingly, it is becoming an influential factor in the level of customer brand loyalty and market perception. As more and more businesses and organizations appreciate the role of good customer support in their service offerings however, the more the definition of "good customer support" changes relative to incresing consumer expectations and requirments. Lately, speed, personlization, intuitiveness and automation are some of the key factors that are shaping customer perceptions of a good or bad customer support experience.

1.2 PROBLEM STATEMENT

As earlier stated, the quality of customer support rendered is fast becoming one of the most influential factors for customer retention in an increasingly competitive business environment in Uganda, let alone the world at large. However, despite this realization by various businesses, both the absense and the high costs associated with administering quality customer support have hindered their efforts to meet user expectation in this regard. For instance, probaby because of the high costs associated with hiring and maintaining full-time customer support personnel and teams, coupled with the some of the highly human-dependent customer support management systems in use today, cases of unresponsive customer support channels have been commonly reported by customers [Carlos, 2019, Abola, 2016]. In cases where the channels are responsive, for instance in automated telephone systems usually used in the telecom and public service sectors, over 56% of consumers have said that themost frustrating thing abount that kind of customer support is the that it makes it hard for them to reach a human agent [Lin, 2022].

To the businesses that receive the customer support complaints, routing and escalating these complaints among support representatives and the technical personnel that can handle them better has proven to be not only rather challenging, but also monotonous in nature. Consquently, this has left some customer support teams, mostly in small and medium sized entreprises with limited resources, not only fatigued but also unmotivated to work at their best levels. They then tend to have very low complaint resolution rates and very high issue resolution durations. Generally, the current state of customer support in Uganda is in such a dire state that Mr. Kalyegera Timothy, a renowned social and political researcher, in one of his articles that addresses the matter, analogously stated that more Ugandans suffer daily at the hands of the poor customer care service at their mobile phone companies, hospitals, passport office, Uganda Revenue Authority branches, banks, schools, restaurants and airline check-in counters than suffer from rigged elections [Kalyegira, 2021], I statement they I could personally agree with to a certain degree.

1.3 RESEARCH OBJECTIVES AND QUESTIONS

1.3.1 MAIN OBJECTIVE

This study seeks to propose machine learning algorithms for automating customer support in the telecom industry in Uganda.

1.3.2 SPECIFIC OBJECTIVES

This study specifically seeks to;

1. Explore twitter as a customer support channel in the telecom industry.
2. Determine the limitations of twitter as a customer support channel.
3. Propose a suitable solution based on a machine learning algorithm (model).
4. Analyse the effectiveness of this machine learning model in this regard.

1.3.3 RESEARCH QUESTIONS

This study seeks to conclusively answer the following research questions;

1. What are the consumer perceptions on the existing customer support methods?
2. How effective is Twitter as a channel for offering customer support?
3. Which processes in the customer support interaction can be automated?
4. How best can machine learning be leveraged to implement this automation?

1.4 RESEARCH SCOPE

1.4.1 GEOGRAPHICAL SCOPE

Geographically, this study will be scoped to Uganda [1.3733°N,32.2903°E], a landlocked country in East Africa. We shall base our research on the telecom in industry in Uganda, with a primary focus on its two largest telecom service providers, that is — MTN Uganda and AIRTEL Uganda, which collectively account for over 90% of the share of mobile subscribers in Uganda, as of the third quarter of 2022 [Statista, 2021]. It is to note, that Twitter will be highly leveraged in this study, and the fact that its jurisdiction of operation is global should be acknowledged, in this regard.

1.4.2 TECHNICAL SCOPE

This research study seeks to investigate the current customer support offered through Twitter in the telecom industry. Its coverage will consist of both a qualitative investigation of the current state of customer support through an online survey posed to 155 random volunteers, as well as a quantitative analysis of the efficiency of twitter as a channel for customer support in the telecom industry, which will be achieved by analyzing 1003 customer support tweets sent to both MTN and AIRTEL on Twitter.

The main key performance indicator that will be benchmarked to ascertain the efficiency of a telecom service provider at handling customer support tickets sent through Twitter will be the "First Response Time". This will be the difference between the creation of a customer support tweet and the first tweet response from the corresponding telecom company on Twitter. Hence, it is to highlight, that this study will not explicitly investigate how these companies handle the customer support tickets sent through Twitter at their premises or trouble ticket systems, nor will a physical correspondence with these companies be required. Also, it is to note, that the end-product of this study is not a trouble ticket system but rather machine learning models that can be integrated in any custom trouble ticket system to automate a certain subprocess.

1.5 TIME SCOPE

This research study will officially run from 1st September 2022 untill the 11th October 2022. The tweets that will be analyzed will be not have been created at any date after the 15th Septemember 2022. Otheriwse, the final say on the time frames of this study adhere to regulation from the project supervisor, Mr. Kizito Jonathan, and we, the researchers, are bound to operate within them.

1.6 RESEARCH SIGNIFICANCE

Since the general state of customer support in Uganda is in a dere state, this research study could primarily serve as a benchmarking document for other academicians, researchers and entrepreneurs that seek to innovate as well as implemenent solutions around customer support, most of all, while leveraging machine learning to automate various subprocesses in its administration. Despite having a focus on customer support in the telecom industry, the overall topic of the study — customer support — encompasses many domains. Hence this document could also be adopted by various companies and entreprises that seek to improve their customer support services, especially through Twitter.

In addition to that, the results of this study could be used to inform and guide regulatory bodies in the telecom sector, and in commerce in general, during design and formulation of policies geared towards improving the state of doing businesses by all stakeholders in Uganda. The sixth consumer right concerning redress as stated on the website of the Uganda Communications Commission states that, ”a consumer has a right to an effective sytsem for handling of complaints” [UCC, 2022]. This document can serve to investigate the effectiveness of such regulations.

2 LITERATURE REVIEW

2.1 DEFINITION OF THE CUSTOMER TROUBLE TICKET

A customer trouble ticket, also commonly known as a customer support ticket or issue is a systematic documentation of the interaction between a customer and a service representative at an organization. It includes vital information for the account involved and the isssue encountered - for instance, its source, date and time of occurence, customer details and most of all, the nature of the issue encountered by the customer [Leen, 2022].

In an organizational setting, customer trouble tickets usually fall under four categories, that is; the service request ticket - that usually contains an inquiry for information about a product/service or related request, the incident ticket - that is usually a report of an unplanned interruption or reduction in the quality of service, the problem ticket - that is usually the root cause of an incident ticket, and the change request - that is a request for a change, modification or replacement of a product or service [Singh, 2021].

Ticket Category	**Summarized Description**
Service Request Ticket	Inquiry for product or service information
Incident Report Ticket	Report of unplanned service-quality interruption
Problem Report Ticket	Report of problem with product usage or service status
Change Request Ticket	Request for service modification and feature suggestions

Table 1: Summary of the categories of trouble ticket systems

Furthermore, a trouble ticket can be triggered by either of two forms, i.e., it can be an anticipatory ticket — where customer support is proactively innitiated, foreseeing the support needs of the would-be creator of the ticket at various points during their lifecycle, or it can be a responsive ticket — where customer support is only innitiated post ticket creation, and cannot prevent issues before they crop up [hiverHQ, 2022].

2.2 DEFINITION OF THE TROUBLE TICKET SYSTEM

A trouble ticket system - also known as an issue tracking-, support ticket-, request management-, or incident ticket system is a computer software package that manages and maintains customer trouble tickets at an organization or business [Bertram, 2009]. In an institutional setting, trouble ticket systems are commonly used in an organization's customer support call center to create, update, and resolve reported customer trouble tickets from various channels, or even support tickets reported within the organization itself by its employees.

There is considerable evidence of a number of trouble ticketing systems. The more commonly noted ones evidenced in the extant literature encompass: the helpdesk, service desk, and the customer relationship management system (CRM). In this study, however, we put our focus on the help desk. Thus, according to Wikipedia, a helpdesk is a computerized trouble ticket system that provides the end user with technical resources, information and assistance related to a company's product and services, as a corporate gesture of customer support. Examples of trouble ticket systems on the market today include and are not limited to Zoho Desk, Zendesk, Freshdesk and Jira Service Desk.

2.3 THE LIFECYCLE OF A CUSTOMER TROUBLE TICKET

2.3.1 TRANSITION STATES OF A TROUBLE TICKET

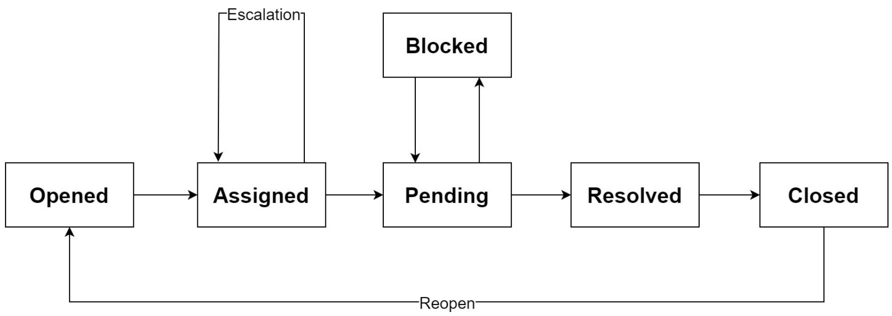

Figure 1: States of a customer trouble ticket in a trouble ticket system

As illustrated in figure 1 above, the average trouble ticket iteratively goes through various stages throughout its journey in a trouble ticket system. At any point in time, it could exist in either of six states, that is; open state — here, the ticket has been created by a client but not yet assigned to a particular support agent, assigned state — here, the ticket has been assigned to a particular support agent(s), pending state — at this stage, interaction between a support agent and a client is ongoing but a final resolution has not yet been found, blocked state — here, either the support agent is awaiting an unreadily available resource that is vital for the resolution of the ticket or an unexpected issue from the internal support department is preventing the resolution of the ticket, resolved state — at this stage, a solution to the customer's issue has been found, shared with them, and implemented, and the customer has gone forward to mark the ticket as resolved, closed state — here, a resolved ticker is marked as closed, and all interactions on it cease.

2.3.2 TRANSITION PROCESSES OF A TROUBLE TICKET

Processes are responsible for the transition of a ticket from one state to another in a trouble ticket system. The efficiency of these processes dictates the qaulity of service of a trouble ticket sytem to its user groups, and presents the basis of our proposition to automate subprocesses in a trouble ticket system as a means to improving the quality of customer support rendered through it. These processes include and may not be limited to;

Ticket Creation

This process entails how a customer creates a support ticket. This can be a guided step-by-step process, depending on whether the customer suuport system is omnichannel or unichannel. An omnichannel trouble ticket system supports the creation of tickets from external platforms like social media platforms such as Facebook, Twitter, and Email, that have no functional connection with it, whereas a unichannel trouble ticket system usually consists of a dedicated interface for collection of customer issues. This could be in form of a dedicated website, mobile application, or a dedicated toll free line, among others.

Ticket Classification

Upon successful creation, an open ticket has both system-generated and user-generated fields — for example; a ticket subject, description, creation date and time, manually or automatically determined ticket category, priority and severity, submitter, among others [Paramesh et al., 2018]. Henceforth, ticket classification, also known as ticket categorization or tagging, refers to the assignnment of a particular ticket to a subset of predefined categories based on the values of its attributes [Sasaki and Kita, 1998].

Ticket Triaging and Routing

Ticket triaging entails the reception of a ticket in a trouble ticket system, tagging it and ascertaining its priority level and then routing it to the appropriate support agent. Traditionally, this process is manual and highly agent-dependent. The efficiency of this process is so important, since it impacts various key performance metrics of the trouble ticket system, as a whole. A misrouted ticket in poorly designed triaging system may take a customer from agent to agent in an endless quest for the right person or department , which is a sure-fire way of decreasing customer satisfaction [Nord, 2022].

Ticket Escalation

Ticket escalation is the process through which a company or organization follows to move a customer issue to a higher-level support agent [Barnecut, 2022]. In systems that heavily rely on manual ticket classification by a customer, which is prone to errors, a ticket can be assigned to an agent or department, that does not possess the necessary resources to resolve it, which necessites its escalation to another higher-level agent or department.

No.	Tier Level	Level Description	Telecom Scenario Ticket
0	Tier-0	Self-service support	How do I check my account balance?
1	Tier-1	Basic front line support.	May you please reset my loan limit.
2	Tier-2	Advanced front line support	I can't send or receive mobile money?
3	Tier-3	Topic expert support	My mobile money account is blocked.
4	Tier-4	External support	I've not received my Yaka token.

Table 2: Summary of the five levels of customer support.

As illustrated in table 2 above, there are five levels of customer support that can be offered, namely; Tier-0 — that consists of all the tools that a company puts at the customer's disposal to help them fix the incident themselves, such as self-service portals, service catalogs, knowledge bases, blog posts and chatbots, Tier-1 — where personnel and support teams start to get directly involved in basic technical support issues, and where most of the ticket escalation on a particular ticket is likely to occure, Tier-2 — where more technically knowledgeable support personnel provide in-depth trouble shooting and backend support, Tier-3 — the highest level in terms of internal customer support and the support personnel at this level have access to the highest levels of technical resources and are usually subject-matter experts, Tier-4 — this is all about outside technical support provided but not supported by the company [Mancilla, 2022].

Ticket Problem Identification

Also referred to as ticket problem mining, this is the extraction of the knowledge of the actual poblem from within the ticket description [Shimpi et al., 2014]. Traditionally, this is manually executed by the agent assigned to the ticket, by reading the ticket description and trying to pick the actual technical issue.

Ticket Solution Identification

This is the selection of an appropriate solution from a solutions database or knowledge space, in response to a particular problem derived from a support ticket description. Likewise, as is the case in ticket problem identification, this process is largely manual in most traditional organizational settings, despite being the central and defining process in the efficiency of a customer support team, let alone system.

Ticket Reopen

A ticket reopen happens when a customer replies to a closed ticket or manually changes its status from closed to open through the customer support portal. This process is usually an indicator of poor perfomance in one or more of the other prior processes in a trouble ticket system, and it is in the best interests of a good customer support experience, that it is tracked and mitigated.

2.4 CUSTOMER SUPPORT KEY PERFORMANCE INDICATORS

A key peformance indicator (KPI) is the metric of collecting, analyzing and/or reporting information regarding the performance of an individual, group, organization, system or component. Also largely known as business process management, the measurement of key performance indicators in an organization subsetting serves many purposes: planning, evaluation, organization learning, driving improvement efforts, deciossion making, resource allocation, control, facilitating the devolution of authority to lower levels of the hierarchy, and helping to promote accountability [Kravchuk and Schack, 1996].

One proverb can be found in literature: "if you want to improve something, measure it" [Radovic and Karapandzic, 2005]. It is from this perspective, that as we propose full-scale automation as a means of improving the perfomance and efficiency of traditional web-based trouble ticket systems, we capture and account for the key perfomance indicators that will prove the feasibility and effectiveness of this proposal. Below, we give a brief account of the 11 most-crucial key performance indicators in customer support, that should be highlighted in the service-level agreement (SLA).

Customer Satisfaction Score (SCAT)

Customer satisfaction is a measure of customer sentiment, or rather how a company's customer support meets or surpasses customer expections or their specified satisfaction goals [Farris et al., 2010]. It is tracked through either posing a simple question to the customer post-support, such as: *On a scale of 1 — 10, how satisfied are you with your recent customer support experience?* or through displaying a star rating system. Based on customer responses, a customer satisfaction score for a company's customer support can be calculated using the formular below, with a score above 80% considered optimal:

$$\text{Customer Satisfaction, CSAT (\%)} = \frac{\text{Total Response Scores Given}}{\text{Total Possible Response Scores}} \times 100$$

Customer Effort Score (CES)

Customer Effort Score is a single-item metric that measures how much effort a customer has to exert to get an issue resolved, a request fulfilled, a product purchased/returned or a question answered [Farris et al., 2010]. It is expected that the effort required of a customer to have their issue resolved should in someway proportionally vary with the technicality of their trouble issue. It is tracked by posing a question or statement like; *To what extent do you agree with the following statement: "[Placeholder for company name] made it easy for me to handle my issue."*, and the respondents can choose from for instance seven answer choices ranging from *strongly disagree* (Score 1) to *strongly agree* (Score 7). Based on customer responses, the customer effort score can be calculated using the formular below, with a score of 5 and above in this case being optimal:

$$\text{Customer Effort Score (CES)} = \frac{\text{Sum of Highest Possible Responses}}{\text{Sum of Received Responses}}$$

Employee Satisfaction Score (ESAT)

Employee satisfaction is a measure of how happy a company's internal customer support agents are with their job and working environment [Bhatti and Qureshi, 2007] — a percentage of customer support agents that have above-average job satisfaction. In our particular context, we focus on the causal factors behind the correlation between the agents' average satisfaction score and the user experience of the trouble ticket system in use. This can be tracked and measured using either of four methods; an employee satisfaction survey, using the employee satisfaction index (ESI), using the employee net promoter score (eNPS) or having 1-on-1 meetings.

In the eNPS, a question such as, *On a scale of zero to ten, how likely are you to recommend our product to another external support agent?*, is posed to the agents, and based on their answers, they agents can be classified as; $eNPS = \%$ of promoters $- \%$ of detractors.

Employee category	Detractors							Passives		Promoters	
Promoter Score	0	1	2	3	4	5	6	7	8	9	10

Table 3: Categories of employees in a eNPS index.

Total Ticket Volumes

This is entails tallying the total number of tickets opened across various dimensions in the trouble ticket system. Some of these dimensions may include and not be limited to; ticket volumes per customer, per agent, per channel, per date and duration, per topic, per category, per status, among others.

First Response Time (FRT)

Also referred to as First Reply Time, FRT is a measure of the time elapsed between a customer raising a ticket and agent's innitial response to it. Faster response times demonstrate to customers that you are here and ready to help them, even if that first response is a brief reply to reassure them that you're looking into the issue. It was found that 96% of customers expect at least a response — as abare minimum — within the 24 hour timeframe, and regardless of the channel. [Ward, 2019]. It can be calculated using the formular below;

$$\text{Average First Response Time} = \frac{\text{Sum of First Response Times of Selected Tickets}}{\text{Number of Tickets}}$$

First Contact Resolution (FCR)

Also known as First Call Resolution, FCR is a metric used to measure customer inquiries or problems resolved on the first call or contact with a respresentative or agent — the percentage of innitial support tickets that do not require further contact after the first. It can be calculated using the formular below;

$$\text{First Contact Resolution, FCR (\%)} = \frac{\text{Total Tickets Resolved on First Contact} - \text{Total Tickets Reopened}}{\text{Total Tickets}} \times 100$$

Average Resolution Time (ART)

Also known as the Average Handle Time (AHT), ART is the average time it takes an agent to resolve a customer service query. In our particular context, we focus on the causal factors behind the correlation of average handle time and efficiency of the various processes in existing trouble semi-automated trouble ticket systems vis-à-vis our proposed system. It can be calculated using the formular below, with the optimal ART in the telcom industry being roughly 9 minutes [Bennett et al., 2020].

$$\text{Average Resolution Time (ART)} = \frac{\text{Total End-To-End Resolution Duration of Tickets}}{\text{Total Tickets}}$$

Cost Per Resolution (CPR)

Also known as Cost Per Ticket (CPT), this is an objective measure of the total monthly operating expense of employing a trouble ticket system dvided by the number of support tickets received in that month. It can be affected by various factors such as level of agent utilization or automation, ticket handle time, wage rates, turnover and absenteeism, among others. CPR can be calculated using the formular below;

$$\text{Cost Per Resolution (CPR)} = \frac{\text{Total Operational Cost of Trouble Ticket System}}{\text{Total Ticket Resolved}}$$

It is to note, that the total operational cost of a trouble ticket system should include and not be limited to the salaries and benefits of the help desk agents, technology and telecommunication expenses, software licencing fees, operational equipment, facilities expenses, utilities, insurance, related travel, and training [BMC, 2022].

Net Promoter Score (NPS)

In a help-desk, net promoter score is a metric intended to measure customer loyalty. Similar to the eNPS in sub section 2.4 on page 8, it tracked by sending out an NPS question, such as *On a scale of 1 — 10, how likely are you to recommend Happyclient to a friend or colleague?*, and categorizing customer responses under promoters passives and detractors. The net promoter score can then be calculated using the formular below;

$$\text{Net Promoter Score} = \text{\% of Promoters} - \text{\% of Detractors}$$

Customer Retention

This is a measure of the ability of a customer support team to prevent customer attrition — the loss of clients or customers, as a result of poor customer support or experience. In an a trouble ticket system optimized to track and measure it, it can be calculated using the formular below;

$$\text{Customer Retention (\%)} = \left(\frac{\text{Total Customers Now} - \text{Customers Acquired During Period In Focus}}{\text{Total Customers Then}}\right) \times 100$$

2.5 CUSTOMER SUPPORT IN THE TELECOM INDUSTRY

2.5.1 KEY CHANNELS OF CUSTOMER SUPPORT IN-USE

Social Media

According to Wikipedia, social media (social networks) are described as interactive technologies that facilitate the creation and sharing of information, ideas, interest, and other forms of expression through virtual communities and networks. User-generated public content websites or application such as YouTube, Twitter, WhatsApp and Facebook are offering avenues for complaining today [Tripp and Grégoire, 2011].

Interactive Voice Response Systems (IVRS)

Usially accessed behind a toll-free telephone line, IVRS is a technology that allows humans to interact with a computer-operated phone system through the use of voice and inputs via a keypad, prompting the IVRS to respond with pre-recorded or dynamically generated audio to further direct users on how to proceed. For instance, in Uganda, MTN and AIRTEL subscribers can access their IVRS by dialling *100* respectively.

Emails

Immediately after the advent of the internet and its world-wide adoption in the 1990s, emails were positioned as the de-facto customer support channel for most big entreprises, besides the call center. Fast forward to today, and this has barely changed. Almost evry organization has an email dedicated to receiving consumer complaints as well as sending out consumer support at scale. These email channels are usually connected to a trouble ticket system through which all these interactions are managed, monitored and analyzed. For instance, the customer support emails for both MTN and AIRTEL are customerservice.ug@mtn.com and customerservice@ug.airtel.com respectvely.

Web-based resources [FAQs, Forums, Self-Help, Chatbots]

Many companies and organizations are using the web to offer minimal-human-dependent customer support using various web technologies today. Rarely updated Frequently Asked Questions (FAQs) are often embedded in company websites, and Forums are used where cuustomers interact amongst each other, usually to put forward feature suggestions or reply to questions from fellow customers that might not require agent intervention. In addition to that, customers have been referred to self-help repositories of knowledge and for the few companies that have attempted to adopt automation, third-party chatbots have been embedded in their websites which usually route customer queries to other resources or agents.

2.5.2 CASE-STUDY: TWITTER AS A TROUBLE TICKET SYSTEM CHANNEL

Businesses have been using Twitter for marketing purposes since its inception, but over the last few years they have increasingly been using it for customer service too [Digivate, 2022]. In some rather rare case, a few create separate handles from their official brand handles, specifically dedicated to live customer support. Social media customer service stats show that 64% of customers on Twitter say they would rather message a dedicated support handle than call a business [Twitter, 2022b].

Figure 2: A customer support ticket to Airtel on Twitter

2.5.3 MERITS OF TWITTER AS A TROUBLE TICKET SYSTEM CHANNEL

Flexibility to the customer

In an omnichannel trouble ticket system, social media platforms, Twitter inclusive, are probably one of the most flexible platforms to collect complaints from customers, and have some of the lowest customer effort scores (CES) (see section 2.4). In contrast to for instance sending an email, where the customer needs to explicitly know the email address for customer support, sending a support ticket is as easy as searching the business on Twitter and either tweeing them or sending a direct message.

Faster responses

Compared to support email that usually go unanswered for days, the engagement of brands with their customers as regards customer support is usually faster and seamless, since it overlaps with their marketing environment.

Affordable alternative

To very many small- and some mid-sized businesses and entreprises, the management of their customer trouble tickets at scale is most cost-effective when implemented through social media platforms like Twitter. To be more specific, since most trouble ticket systems in use made by international companies that pricing plans on international standards, most companies simplycan't afford the average monthly cost per seat of such systems, that stands at approximately 10s of dollars.

2.5.4 DEMERITS OF TWITTER AS A TROUBLE TICKET SYSTEM CHANNEL

Character limit

A tweet can have upto 280 characters [Twitter, 2022a]. This puts a limitation on the expressibility of both customers and support agents. In the long run, fewwer and fewer customers find it as aviable option to report their issues, most especially complex ones.

Exposure of sensitive information

Naive consumers that are faced with a challenge will not hesitate to reach out to customer support through a tweet, usually exposing sensitive information such as mobile numbers, emails and names, location details, device details among others. This posses a security risk not only to them, but in some instances, to the companies too.

Volume of Content

Before being a medium for administering customer support, Twitter is a social media platfrom for social conversations. The intersection of customer support and this social conversatory nature leads to a high volume of complaints, where some are conversatory and not necessarily requiring any form of support or rather action, but only acknowledgment. This usually leads to high ticket backlogs and necessitate automation.

Negative exposure

From the perspective of businesses, the creation and resolution of tickets publicly usually has a negative impact on their brand and consumer sentiment. This is the reason most businesses opt for emails and the call center as their first choice for receiving and resolving customer support complaints.

Uncustomized functionality

Because Twitter is a third party channel for the collection of customer complaints, companies usually have less power over customizing their customer support experiences. For example, tweet lengths are limited, image quality and quantity is highly limited, and companies have less access to user details, which are necessary when offering customer support. For example, it was noted that the first reply to a customer complaint on Twitter from a telecom company was, What is your phone number and location, please?, a question that would not be necesssary in a dedicated customized and personlized customer support environment.

2.6 TICKET CLASSIFICATION THROUGH SENTIMENT ANALYSIS

2.6.1 LITERATURE REVIEW ON SENTIMENT ANALYSIS

Sentiment Analysis has many names. It's often referred to as subjectivity analysis, opinion mining, and appraisal extraction, with some connections to affective computing (computer recognition and expression of emotion) [Pang et al., 2008]. One of its main challenge is to define the objects of the study — opinions and subjectivity. As a field of research, it is closely related to (or can be considered a part of) computational linguistics, natural language processing, and text mining [Mejova, 2009].

Sentiment that appears in text comes in two flavors: explicit where the subjective sentence directly expresses an opinion ("It's a beautiful day"), and implicit where the text implies an opinion ("The earphone broke in two days") [Jindal and Liu, 2006]. Most of the work done so far focuses on the first kind of sentiment, since it is the easier one to analyze.

Another important aspect in sentiment analysis is the polarity feature of the study text. It is usually dichotomised into two– positive and negative– but polarity can also be thought of as a range. Furthermore, a distinction must be made between the polarity of sentiment and of its strength [Pang et al., 2008]. One may feel strongly about a service being OK, not particularly good or bad; or weakly about a product or service being very good (because perharps one has used it for a very short time to form a stronng opinion).

Another important part of sentiment is its target- an object, a concept, a person, anything. Most work has been done on product and movie reviews, where it is easy to identify the topic of the text. But it is often useful to pay attention to which feature of this object the writer is talking about: is it the camera display or battery life that troubles consumers the most? Because of ready availability of product review datasets, feature extraction has been closely studied in the past decade [Jindal and Liu, 2006]). The mention of these features in text can also be explicit ("Battery life is too short") or implicit ("Camera is too large") [Jindal and Liu, 2006].

2.6.2 MODE OF APPLICATION

The ticket triaging and routing processes (refer to section 2.3.2) in a trouble ticket system presents the best opportunity for automation. Traditionally, tickets are not only serviced on a first-come-first-serve basis but also predominatly manually tagged and routed, a rather inneficient method for scaled customer trouble ticket systems, such as in the telecom sector. Both mood-based and aspect-based sentiment analysis can be leveraged in this scenario to offer an individualized service to each ticket. The sentiment of a ticket can be used to automatically predetermine its priority level. On the other hand, the aspects found in a ticket can be used it determine service pipeline, or rather to which support agent or team it should be routed to as well as come up with classifications or ticket groups that would require similar service.

In this particular use-case, two libraries were used. The first one isPyABSA, an opensource aspect-based sentiment analysis (ABSA) framework that includes features of aspect term extraction, aspect sentiment classification, and text classification. Furthermore, multiple ABSA subtasks can be adapted to PYABSA owing to its modular architecture. Out of the box, it comes trained on over 21 ABSA datsets in over 8 languages that cover several domains [Yang and Li, 2022]. This library was used to extract the aspects from each ticket as well the corresponding sentiment and confidence levels.

The other library was the TextBlob, a Python (2 and 3) library for processing textual data. It provides a simple API for diving into common natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more [Loria et al., 2018].

Given a text of not more than 80 words, post-process, pyABSA returns an object with the following results; an array containing the constituent words of the input sentence(s) — bag of words, an array that can be either emtpy or containing the aspects found in the input sentence(s), an array containing the corresponding discrete sentiments (Positive, Negative, or Neutral) of the aspects found, and lastly an array containing the corresponding levels of confidence of the model towards each of the aspects. TextBlob on the other hand returns the polarity, subjectivity of the input text. The figure below shows the combined results from both models given the input text, *"I hate their customer support although I love their internet speed."*

Input Text	"I hate their customer support although I love their internet speed.
Text Subjectivity	0.75
Text Polarity	-0.15000000000000002
Aspects	["Customer support", "Internet speed"]
Sentiment Per Aspect	["Negative","Positive"]
Confidence Per Aspect	[0.9998087286949158, 0.9998565912246704]

Table 4: Sample results of PyABSA and TextBlob

3 RESEARCH METHODOLOGY

3.1 INTRODUCTION

In this chapter, we present and justify the methodology and methods employed during the data collection and analysis phases for this study. Specifically, we account for the choice of participants, the criteria for their inclusion in the study, as well as how they were sampled. A description and justifaction of the research design used in the study is presented, on top of stating the instruments, their reliability and validity and the procedures employed during data collection. We conclude by discussing how the data was analyzed and put forward the ethical issues encountered during the study.

3.2 RESEARCH DESIGN

This research study will follow a descriptive research design. We seek to explore the viability of automating customer support in the Ugandan telecom industry using machine learning algorithms. The study will follow a more quantitative approach to data collection and analysis. A guided survey will be sent out to a random sample of a chosen population to ascertain their opinions perceptions on various aspects of existing customer support, specifically through Twitter, in the telecom industry in Uganda. In addition to that, public support tickets sent through Twitter will be analyzed to ascertain the merits and demerits of the trouble ticket systems being used to manage them by the recipients. Automatable processes will be assigned and appropriate machine learning solutions will be proposed.

3.3 AREA OF RESEARCH STUDY

The study area is Twitter, a microblogging and social networking platform. It is one of the social media platforms where people make both public and private service complaints to brands and businesses most. We focused on two accounts; the MTN Uganda and the Airtel Uganda accounts, which when combined, their clientele base accounts for a substantial size of the total telecom subscribers in Uganda. As regards to the investigation of user perspectives on existing methods and channels of customer support, our quantitative research was conducted at Makerere University.

3.4 STUDY POPULATION AND SAMPLING TECHNIQUES

The study population comprised of two populations — students at Makerere University between the ages of 18 and 26 and customer support tickets sent through Twitter and are publicly accessible. In the first case, we used simple random sampling — wher a random survey was sent out to Makerere University students and the first 155 responses were taken into account. In the second instance, we used stratified non-probability sampling — where the support tickets where collected under two classifications, with appriximately each 50% of the 1003 tickets to be collected being sent to either of the two telecom companies in our study, that is, MTN and AIRTEL.

3.5 METHODS OF DATA COLLECTION

3.5.1 ONLINE SURVEY

An online survey is a structured questionnaire that a target audience completes over the internet generally through filling out a form. This method enabled us gather primary data on our study objectives in a more scalable manner, compared to other data collection methods like interviews that would otherwise have been not only monetarily draining but also a strain on our financial budget for the study.

3.5.2 WEB SCRAPING

Web scraping is a technique for converting unstructured web data into structured data that can be saved and analyzed in a central spreadsheet or database. This enables the bot to retrieve large volumes of data in short amount of time, which is advantageous in today world especially since we have big data which is always changing and updating [Khder, 2021]. It is usually accomplished using a web scraping framework, with Python as the preferred choice of programming language in most use cases. Data is scraped off a website, cleaned, and stored in either a flat file or a database, for future analysis. This method was chosen since it was the most efficient way of querying primary public data from the internet, within our project time limits.

3.6 INSTRUMENTS OF DATA COLLECTION

3.6.1 GOOGLE FORMS

Google Forms is a cloud-based data management tool used for designing and developing web-based questionnaires.This tool is provided by Google Inc., and freely available on the web for anyone to use to create web-based questionnaires. It was used to create a survey form to investigate the quantitative variables of respondents regarding customer support in Uganda. The form used was titled: ”Investigation of the customer support services to telecom subscribers of MTN and AIRTEL in Uganda”. It consisted of three sections, notably: investigation of demographic distributions, investigation of existing methods, and validation of the need for proposed solutions.

Investigation of demographic distributions

This section outlined the social-demographic characteristics of respondents. Major variables measured here included but are not limited to gender, age range, level of education, primary telecom service provider, and primary social media accounts.

Investigation of existing methods

Here, a thorough quantitative investigation of the respondents was conducted to ascertain the respondents' perceptions on existing customer support channels. These questions sought to further validate and prove the need for automating certain processes in the trouble ticket systems being used in the various trouble ticket sytems of their respective telecom service providers. Questions in this section sought to know their challenges, preferred trouble ticket channels, and satisfaction with the processes involved.

Validation of the need for proposed solutions

Lastly, this section presented multi-choice questions that sought out the essentiality of various propoble solutions to the various demerits of the trouble ticket systems in use by the telecom service providers, and in commerce, generally.

3.6.2 TWITTER API

According to its documentation at *https://developer.twitter.com/en/docs/platform-overview*, this API is described as a set of programmatic end points that can be used to understand or build conversations on Twitter. It allows you to find and retrieve, engage with, or create a variety of different resources including but not limited to the following: tweets, users, spaces, and direct messages, for the Twitter platform.

It is a protected API that requires authentication and authorization to be accessed. We requested for access from Twitter, to which limited permissions were granted to us. We then manually gathered 1000 IDs of tweets (got from their URLs) sent to either the MTN Uganda or Airtel Uganda twitter accounts from and before 15th September 2022. Using the API, we queried the tweet text, creation date and time of tweet, response from telecom company, and date and time of response. The data collected was stored in a csv flat file, to await quantitative data analysis.

3.7 RELIABILITY AND VALIDITY OF RESEARCH INSTRUMENTS

Reliability is the degree to which measures are free from error and therefore yield consistent results (i.e.the consistency of a measurement procedure) [Lakshmi and Mohideen, 2013]. In regards to our study, research items to be measured were subjected to rigorous and robust research procedures in order to produce similar results under similar research conditions. In our particular case, primarily because of our limited time, and monetary resources, a relatively small but particular sample population consisting of university students was deliberately chosen, in order to ensure ease of validation of responses from the respondents.

Validity of the research instruments is the extent to which the items in the instrument measure what they are set out to measure [Miles and Huberman, 1994]. In order to ensure validity of responses in this study, multi-choice questions were used in order to provide respondents with a mental direction towards the objectives of this study, which would not have easily been the case in freely-flowing interview sessions, per say.

3.8 METHODS OF DATA ANALYSIS

Both quantitative and qualitative methods of data analysis were used. Qualitative analysis was majorly conducted on the responses from the survey to ascertain relationships and possible causation factors. Quantitative analysis was primarily conducted on the tweets dataset to ascertain the quantitatively-supporting factors for the automation of the trouble ticket systems that were handling them. The three major Python libraries that were used in this endavour were — Pandas, NumPy and Seaborn — a data visualization library.

3.8.1 DATA WRANGLING

Data wrangling can be defined as the manual or programmatic process of iterative data exploration and trasformation that leads to and enables the analysis of data. The main purpose of this process is to make data usable — put it in a form that can be parsed and manipulated by analysis tools. It consists of sub-processes like data gathering, data cleaning, data enriching, and data validation, among others [Kandel et al., 2011].

First and foremost, structural consistencies in the both the survey responses dataset and the tweets dataset were eliminated. The survey questions were changed to appropriate brief column names. Extensive data cleaning was carried out, where outliers, null values, and duplicates were removed from the tweets dataset, tweet column values were sanitized to eliminate HTML. Here, coding of categorical values was also carried since the python libraries we were intending to use for analysis work best with numerical values.

In addition to that, the raw datasets were enriched, for instance, a new derived column *first reply time* was added to the tweets dataset based on the ticket tweet creation time, and creation time of the response tweet from the respective telecom company. The datsets were validated for both completeness and accuracy so as to ensure correct results were to be derived from our analysis.

3.8.2 EXPLORATORY DATA ANAYLYSIS

According to Wikipedia, exploratory data analysis is termed as an approach of analyzing a dataset to summarize its main characteristics, often using statistical graphics and other data visualization methods. The goal of this step is not to necessarily explain the analysis to the reader but rather for the researcher to gain a deeper understanding of their data. Pandas and NumPy were the two libraries used in this regard, alongside Jupyter Notebook sessions. The shape, descriptive information and statistical description of the datasets were among the few aspects investigated at this stage.

	chatbot_acceptance	main_telecom_cs_csat	rating_of_national_cs
count	155.00000	155.000000	155.000000
mean	3.03871	3.335484	2.993548
std	1.23206	0.962123	0.841279
min	1.00000	1.000000	1.000000
25%	2.00000	3.000000	3.000000
50%	3.00000	3.000000	3.000000
75%	4.00000	4.000000	3.000000
max	5.00000	5.000000	5.000000

Figure 3: Exploratory analysis of a few columns of the survey responses datset

3.8.3 EXPLANATORY DATA ANALYSIS AND VISUALIZATION

Explanatory data analysis is the phase of data analysis that seeks to find answers to our research questions, and usually involves extensive visualizations meant for the reader. At this stage, we picked a subset of the data at hand and used to drive our explanations home in our research result and discussion of results sections below. Here, our Jupyter Notebooks were converted to interactive slides that were hosted online through GitHub, for public consumption. The library that was extensively used in this phase was the Seaborn library, a library used for making visualizations in data analysis using the Python programming language.

3.9 ETHICAL CONSIDERATIONS

Innitially, an endavour was taken to explain the purpose, methods, and goals of the study to the participants.This was clearly highlighted in the survey form shared with them, and participation demanded consent and agreement. Confidentiality and anonymity of the participants was highly considered and implemented and participants were permitted to withdraw their participation at any point in time. Email addresses, names, nor location or contact details of participants were not collected.

Access to the Twitter API had to be requested from Twitter, before any form of scraping was done on the platform. A user policy was shared with us, the researchers, and we endavoured to follow and abide by it during our research study. Upon collecting tweets, all self identifying information that was not vital to our study was removed and the tweets were then aggregatively used. Only publicly assessible information was scraped and nothing was shared with any other third party.

4 RESEARCH RESULTS

4.1 INTRODUCTION

This section looks at the presentation of the results derived from both our qualitative and quantitative analysis of both the dataset of survey responses and the dataset of support tickets scraped off twitter, respectively.

4.2 DEMOGRAPHIC DISTRIBUTION OF RESPONDENTS

Below, an analysis of the profiles of the 155 respondents consisting of young adults between the ages of 18 and 28 is presented. Major emphasis was placed on their gender distribution, the telecom service providers (MTN or AIRTEL) that they primarily subscribe to and the social media platforms on which they have accounts, and regularly use.

GENDER DISTRIBUTION OF RESPONDENTS

The figure on the right shows the gender distribution of respondents. Of the 155 respondents, roughly two thirds of them were male.The online survey that was used to collect the data was circulated in the final year computer science class of Makerere University, 2022 that has a similar gender distribution, which could explain this kind of end result in the gender distribution.

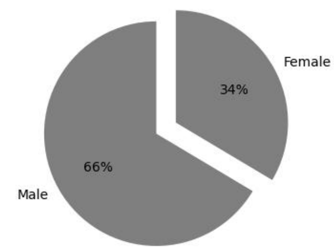

Figure 4: Gender distribution of respondents

DISTRIBUTION OF RESPONDENTS AMONG TELECOM SERVICE PROVIDERS

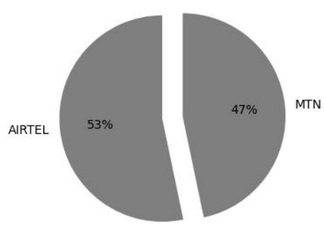

Figure 5: Primary telecom subscriptions of respondents

Of the 155 respondents to the survey, 82 of them considered airtel to be their primary telecom service provider. It is to highlight, that for a respondent considering one telecom service provider primary does not imply that they aren't subscribers of the other. This attribute was crucial to capture because the two telecom companies, that is — mtn and airtel, provide different qualities of customer support to their respective clients.

DISTRIBUTION OF RESPONDENTS AMONG SOCIAL MEDIA CHANNELS

The figure below shows the social media channels on which the respondents have an acount. The two most common social media channels among respondents were WhatsApp (146) and Email (135), followed closely by Twitter (115) — our study case channel, and then Telegram (105) and Facebook (93), in that order.The nation-wide blockage to the Facebook platform by the government could justify why it had the lowest user count. Otherwise, on average, a social media platform had 118 social media users of the respondents.

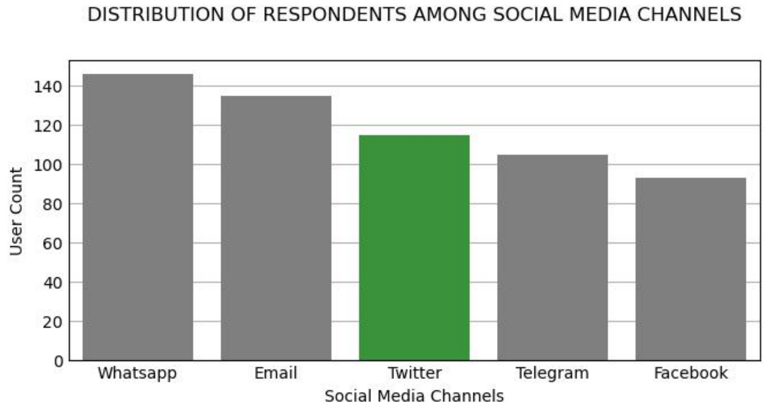

Figure 6: Social media channel accounts owned by respondents

4.3 PREFERENCES OF RESPONDENTS

MOST PREFFERED CHANNELS FOR CUSTOMER SUPPORT

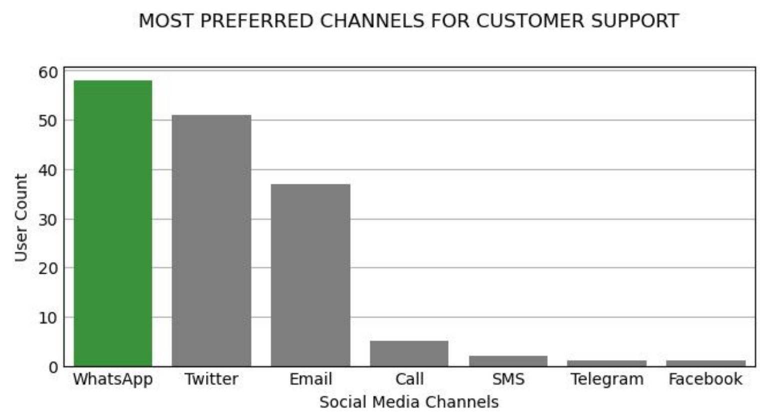

Figure 7: Preffered channels for customer support in the telecom industry by respondents

The figure above shows the most preferred channels for receiving customer support in the telecom industry in Uganda, by the respondents. It can be seen, that the three most preferred channels for customer support are WhatsApp (58), Twitter (51) — our study case channel, and Email (37), which collectively account for 94% of user preferences in this regard. Telephone systems such as IVRS (2.5.1) and toll free telephone lines came in fourth (5), followed by SMS and USSD channels (2), with Telegram (1) and Facebook (1) trailing with only one user considering them their preferred channel for customer support.

SATISFACTION TO CUSTOMER SUPPORT THROUGH TWITTER IN THE TELECOM SECTOR

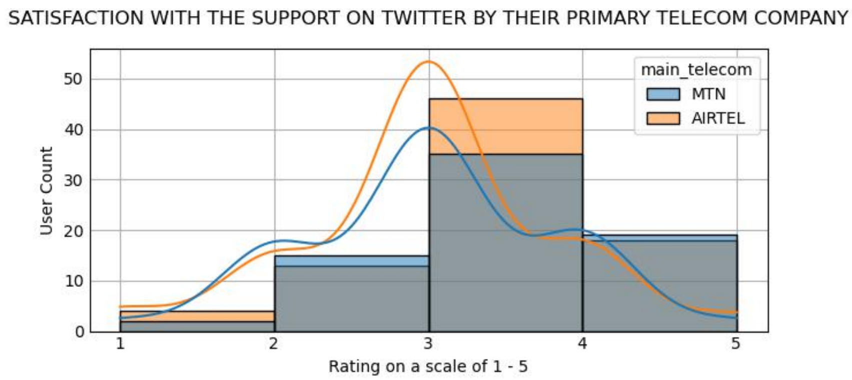

Figure 8: Satisfaction with customer suupport on Twitter by the respondents

The figure above shows the rating of customer support offered by the two primary telecom service providers to their respective clients among the respondents, on a scale of 1 - 5, with one implying the least level satisfaction and 5 meaning the highest level of satisfaction and contentment by the respondent. The average level of contentment was 3 with AIRTEL having higher levels of customer satisfaction compared to MTN, as per the kernel density estimate (KDE) curves.

CHALLENGES DURING CUSTOMER SUPPORT INTERACTIONS

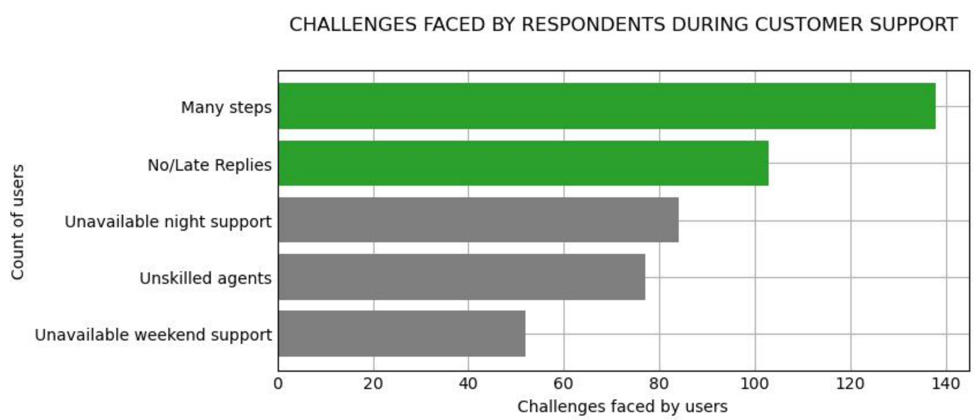

Figure 9: Challenges faced by customers during customer support sessions

The figure above shows the effect of challenges faced by the respondents during customer support sessions. Of the 155 respondents, 138 considered having to go through many steps to finally receive customer support as their biggest challenge during a customer support session. 103 respondents are of the view that receiving late or no replies is a challengeas well. Unvailable support on weekends was not comparably considered a big challenge, perharps due to the fact that some consumers deem it usual not to receive certain services on weekends.

MOST IMPORTANT FACTOR DURING CUSTOMER SUPPORT SESSIONS

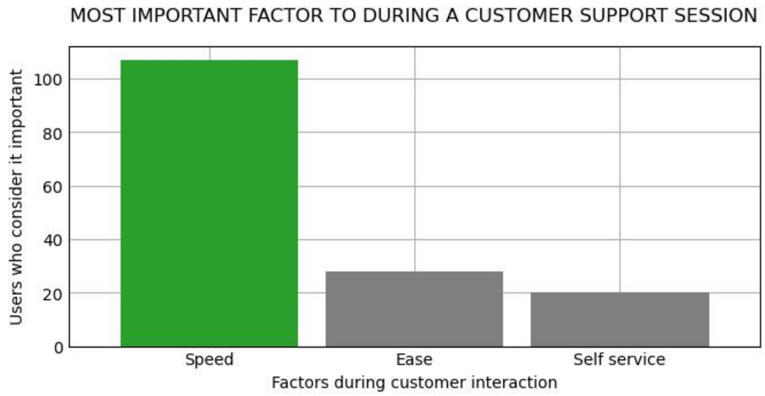

Figure 10: Most important factor during a customer support session

The figure above shows the factors considered important by the respondents during a customer support session. 107, or rather roughly 69% of the respondents consider speed at which customer support is adminstered,be it through an agent or a self service portal, as the most important factor during a customer support session. Ease of getting customer support came in second with 48 respondents and only 20 respondents considered availability of or the ability to receive self-serviced customer support as the most important factor. It should be taken into consideration, that a respondent not considering a factor as most important does not imply that it is not that important at all, but rather that it is that important relative to the available choices.

ACCEPTANCE OF CHATBOTS BY RESPONDENTS

Lately, chatbots are the commonest form of automation that most respondents have probably come across. Here, we attempt to investigate the perception and attitudes of the respondents towards automation as a solution not only during customer support administration, but in the general navigation of the internet. The figure below shows the average acceptance by the respondents to interacting with a chatbot, on a scale of 1 - 5, with 1 implying that the respondent is not willing to interact with a chatbot during a customer support session and 5 meaning that they are very willing to interact with a chatbot when requesting for customer support. The average acceptance rating was 3 with over 50 respondents.

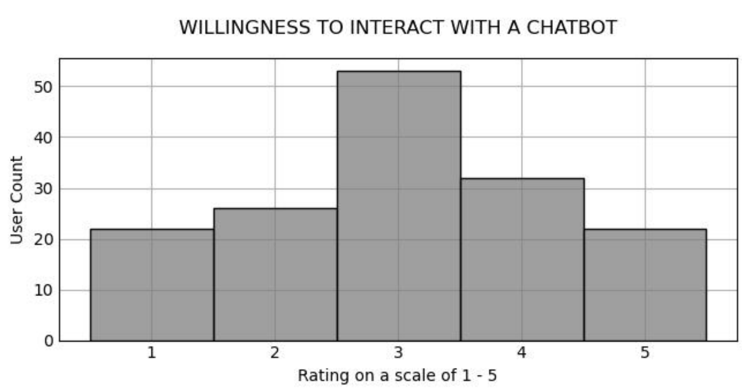

Figure 11: A histogram showing acceptance of chatbots among respondents

MOST ESSENTIAL FEATURES IN A TROUBLE TICKET SYSTEM

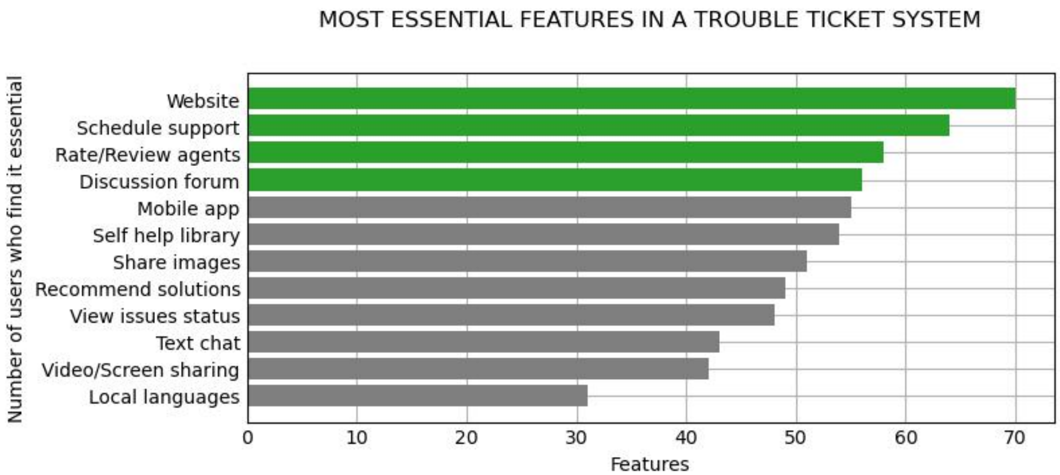

Figure 12: Most essential features in a trouble ticket system

The figure above shows the most essential features ina trouble ticket system, according to the respondents. Most respondents find a website, comparable to a mobile appilication, as the most essential feature. The ability to schedule customer support comes in second, followed by the ability to rate and review support agents post a customer support session. In fourth place is the availability of a forum like space, where users can interact with other users, support agents on aspects that do not require live support. The least essential feature was the availability of multilingual functionalities. However, this could be so simply because our sample population is biased — consisting of only university students and young adults, who predominantly can and usually use english as their primary language.

NECESSITY OF A DEDICATED UNICHANNEL TTS

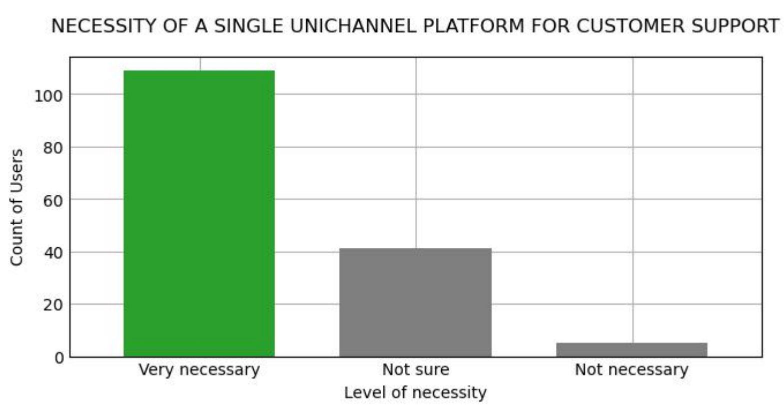

Figure 13: A bar graph of the necessity of a dedicated unichannel customer support platform

Trouble ticket systems can be both omnichannel or unichannel. We investigated the preference of respondents in this regard, since it is easier to automate a unichannel TTS as compared to an omnichannel TTS that are usually integrated with various disparate channels. The figure above shows how necessary respondents would find a dedicated unichannel customer support platform. 109, or rather 70% of the respondents found such a tool necessary, 49% found it somewhat necessary, and only 5 respondents consider such a pltform uncessary.

4.4 ANALYSIS OF KPIs ON TICKETS THROUGH TWITTER

AVERAGE RESPOSE TIME PER TELECOM SERVICE PROVIDER

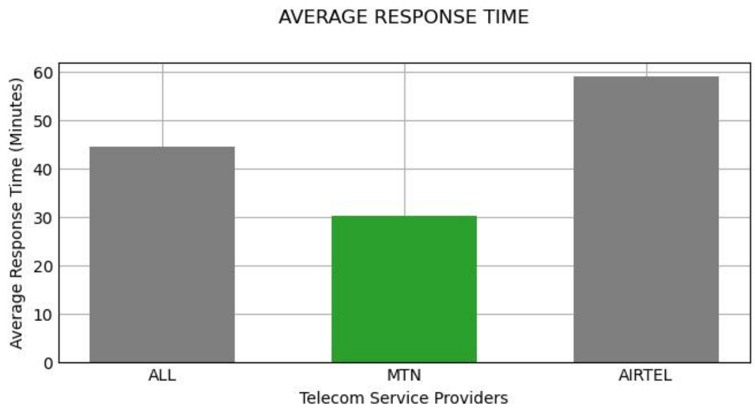

Figure 14: Average Reply Time (ART) of MTN and AIRTEL

The average global average response time (ART) to customer support in the telecom industry is 30 minutes, although this varies per channel.The figure below shows the average response time of the two study case telecom companies to customer support tickets created on Twitter by its clients. The lower the ART of an organization, the more efficient its trouble ticket system is. The average ART was 45 minutes, with MTN having 30 minutes which was way below average and AIRTEL having 59 minutes which is way above average. It is to note, that a low ART could imply high efficiency of a TTS, but not high effectiveness.

TICKETS FIT FOR A FORUM

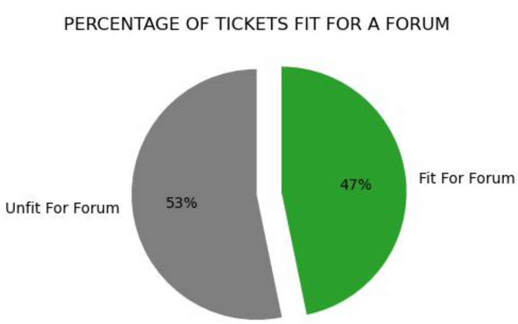

Figure 15: Tickets that were best suited for a forum instead of live twitter support

The figure on the left shows the percentage of tickets that were best suited for a forum-like platform and may not necessarily necessitate live customer agent support but rather response from either fellow customers or an agent at a later time. This aspect had to be investigated because as we propose automation, the highest level of automation is an agent-less experience where customer support is dispensed with minimal human interaction. 47%, or rather 469 of the 1003 tickets that were analyzed were found to have been best sent on a forum instead of a channel that necessitates quick live replies that would otherwise slow response to tickets more suited for live support.

TICKETS FIT FOR A SELF HELP KNOWLEDGE RESPOSITORY

The figure on the right shows the percentage of tickets that were best suited for a self-help knowledge repository where customers could easily find the support they need hence saving live customer support sessions for customers with more techanically demanding requirements. Only 19%, or rather 191 of the 1003 tickets under investigation were found to have been suited for self-help. Likewise, this aspect had to be investigated since the necessity for automation varies with the total number of tickets received by an agent which is likewise influenced by the number of tickets that go through a certain process in the TTS.

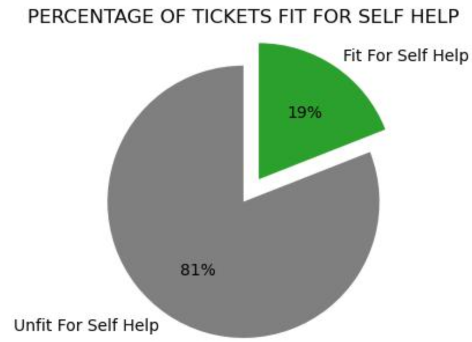

Figure 16: Tickets suited for a self-help knowledge repository instead of live twitter support

DISTRIBUTION OF TICKETS ACROSS TICKET TYPES

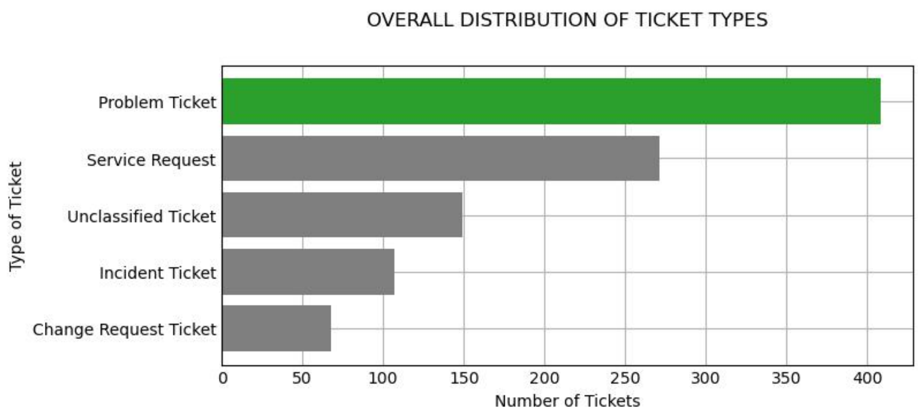

Figure 17: A distribution of the different types of tickets that were analyzed

The figure above shows the different types of tickets that were were sent through Twitter. This aspect was vital enough to track because various automatable processes in a trouble ticket system depend on the ticket type, for instance, ticket triaging and routing, ticket prioritization among others. 41%, or rather 408 tickets were trouble tickets that necessarily need a customer-agent interaction for their resolution. These awere followed closely by service requests (271), where customers demanded or reported information to the company. Unclassified tickets — those that did not fit any form of ticket classification and were merely conversational comments by customers on the respective company accounted for around 15% of the tickets. 107 tickets were incident tickets — those that were simply reporting a general system service interruption but the customer had not yet encountered a problem personaly per say in this regard. Lastly, change requests — those where customers where simply requesting for a feature, a change in operation or functionality were 68 in total.

DISTRIBUTION OF TICKETS ACROSS TIER-LEVELS OF SUPPORT

The figure below shows the distribution of ticket types among the tickets analyzed. Over half the tickets (563) required tier-1 support — basic support from an agent to less technical and unspeacialized tickets. Slighty more than $\frac{1}{3}$ of the tickets required tier-0 support — support that did not require interaction with a support agent but rather self-help material like a self-help knowledge repository, a forum, among others. 73 tickets demanded of tier-2 support, and only 13 and 3 tickets demanded of tier-3 and tier-4 support respectively.

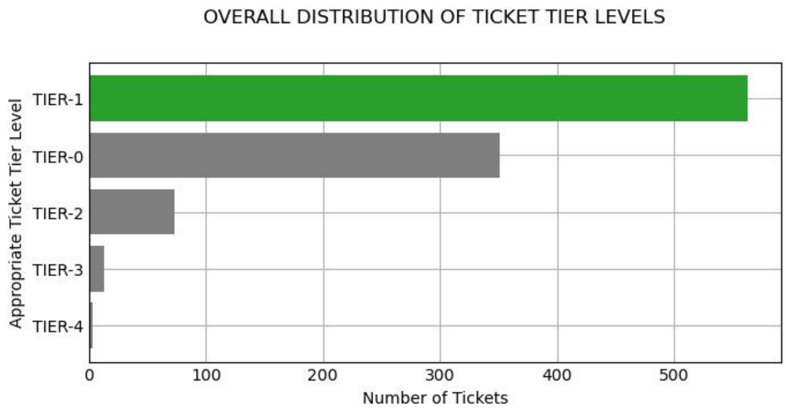

Figure 18: A distribution of the support levels that were required by the analyzed tweets

CORRELATION OF AGENT TICKET COUNT PER DAY AND RESPONSE TIME

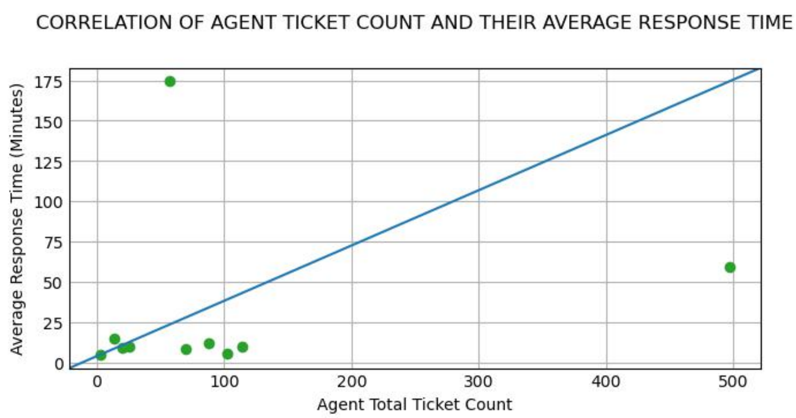

Figure 19: A correlation between an agents total ticket count and their average response time

We sought to investigate the correlation between the total amount of tickets received by an agent in our test-duration with their average response time in that same duration. It is to note, that the result of this investigation were not interpreted as a causal-relationship among the two variable. We used the Pearson correlation coefficient, PCC — the measure of linear correlation between two sets of data. The two variables were found to have a very low positive correlation coefficient (0.1853355947166857), i.e, a positve change in one variable has avery low positive influence in the other variable (an increase in the number of tickets for a perticular agent has a very low incremental effect on their average response time).

4.5 ANALYSIS OF CHOSEN NLP LIBRARIES

4.5.1 EFFICIENCY OF MODELS IN MINING ASPECTS

It took the PyABSA model 655.4137330055237 seconds to analyse and mine the aspects and sentiments from 1003 tickets, averaging 0.65345337288 seconds per ticket. Assuming the average time taken by a customer support agent to read a customer support ticket lies within the range of 30 seconds to 60 seconds, which would imply that the total time taken to manually classify 1003 tickets based on their aspects lies in the range of 30,090 - 60,180 seconds. This is way higher than the time taken by PyABSA.

(a) Wordcloud of raw aspects pre-ABSA

(b) Wordcloud of raw aspects post-ABSA

Figure 20: The reduction in aspects as seen by an agent pre- and post-ABSA

Traditionally, a support agent has to mine the aspects on a ticket-by-ticket basis, despite reocurring aspects in various tickets. However, when a model for aspect-based sentiment analysis such as pyABSA is used, aspects are mined within a very short time and tickets with similar aspects can be grouped together, which can narrow down the range of tickets to look through per aspect by an agent. The figure above shows the first 100 aspects from the 1003 tickets that were analyzed, both before applying an ABSA model on them and after. It can be seen that the number of main aspects (large words) reduces significantly post ABSA.

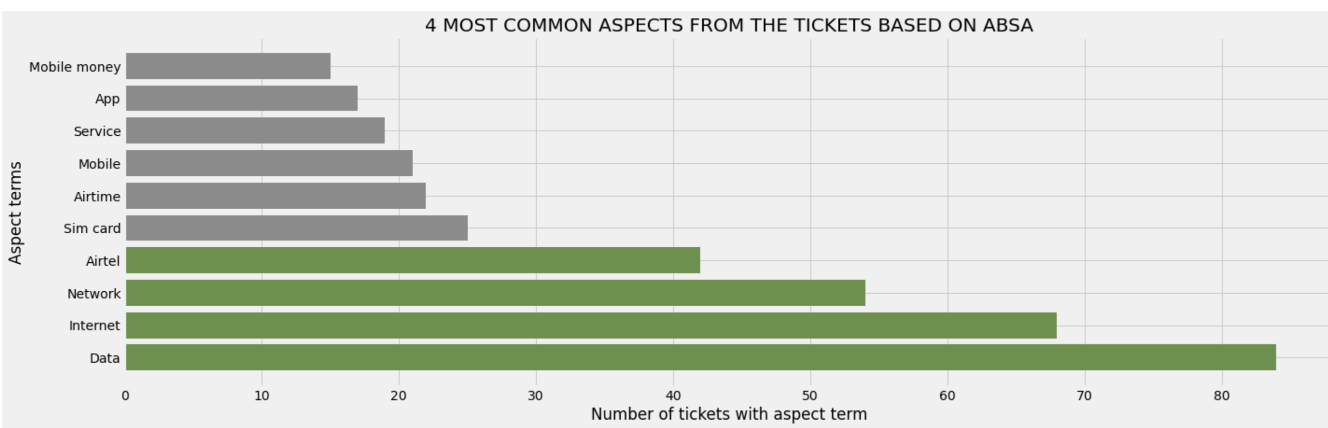

Figure 21: The 4 most common aspects in the 1003 tickets

As seen in the figure above, the four most common aspects where *data* — 84 tickets, *internet* — 68 tickets, *network* — 54 tickets and *Airtel* — 42 tickets, respectively. Because we used a pretrained model, that was trained on an dataset of text from English literature that is not specifically annotated and customized to our resultant test industry — the telecom industry, the dataset fails to find the semantic union between some of the terms. For example, *internet* and *data* are taken as two separate aspects, despite being somewhat similar in the practical context. Also, some aspects are rather brand words that are most likely to appear in many tickets but not necessarily point to a certain product or service per say. For example, Airtel is a brand name, which must have appeared in many tickets, that had other specific aspects to them. However, this perspective should not downgrade the fact that a ticket could have a brand name as its main subject, with its corresponding sentiment attributes.

4.5.2 ANALYSIS OF TICKET- SUBJECTIVITY AND POLARITY

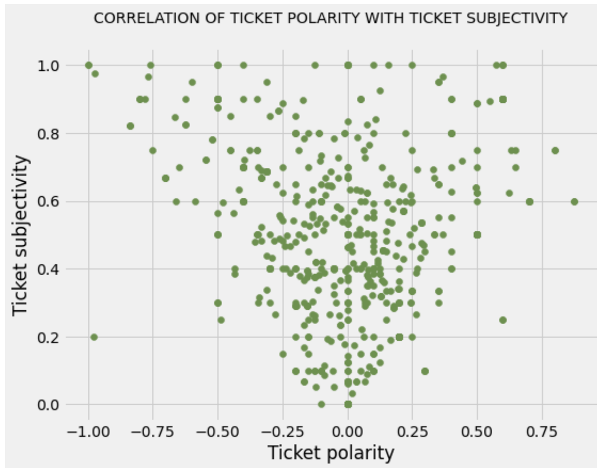

Figure 22: A correlation of ticket subjectivity and ticket polarity

The ticket polarity and ticket subjectivity of a ticket were determined using the TextBlob library. to give a short recap, a ticket's subjectivity quantifies the amount of personal opinion and factual information contained in the ticket. A high subjectivity means that the ticket contains more personal opinion rather than factual information, and it lies within the range of 0 (factual) — 1 (subjective). On the other hand, a ticket's polarity refers to the strength of an opinion in the ticket, and this lies within a rnage of -1 (negative) — 1 (positive).

From the figure above, a few observations were made. Most of the tickets had a subjectivity score below 0.6 with the mean being 0.314687. The lowest subjectivity was 0 and the highest was 1. On the other hand, most of the tickets had a polarity close to zero (neutral) with the average polarity being -0.005123 (negative). The lowest polarity was -1 and the highest was 0.875000. Also, the tickets with lower subjectivity tended to have polarities closer to zero compared to tickets with higher subjectivity.

4.5.3 ACCURACY OF NLP LIBRARIES

TEXTBLOB

The figure below shows the categorization of the tickets by the *textblob* library across the 3 classifications of mood sentiment — positive, negative and neutral. Since these were trouble tickets, most of the tickets were expected to have a negative sentiment. However, upon analyzing them, 45% of the tickets were categorized as ”Positive” and another 45% of them were grouped as ”Neutral”. This implied that only 10% of the tickets had a ”Negative” sentiment. This greatly disagreed with our hypothesis. However, upon traversing literature related to it across the internet, an unverified conclusive statistic was found, — it has an accuracy averaging 65%, which could probably justify this anormaly. Other than that, our datatset could hold many tickets that are classified as positive since most tickets had very low subjectivity.

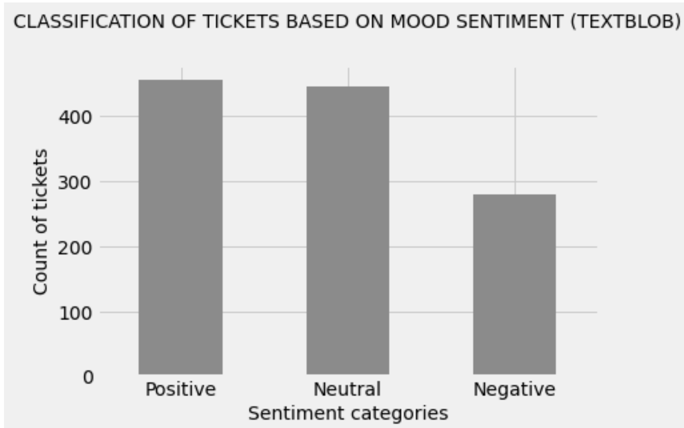

Figure 23: Classification of tickets based on mood sentiment (TEXTBLOB)

PYABSA

The figure on the right shows the categorization of the aspects found in the tickets by the *pyabsa* library across the 3 classifications of mood sentiment. 640 aspects were found to have negative sentiment, 167 were found to have neutral sentiment, and only 57 tickets were found to have a positive sentiment. Based on the nature of the tickets, this distribution agrees with our hypothesis on a higher degree. Notice should be taken that this accuracy is based on a model that was not trained on a dataset with text specific to the telecom industry. Hence, its performance is expected to increase highly once its used on a custom dataset.

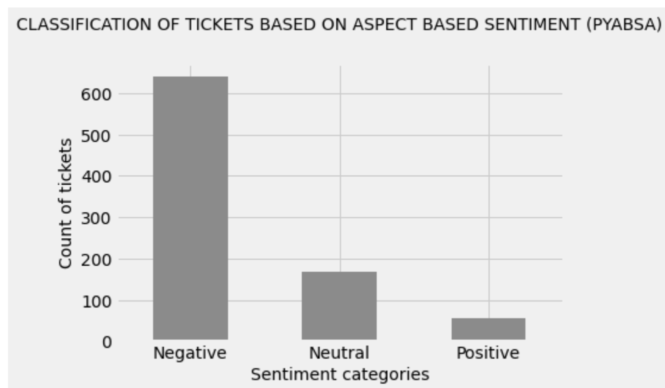

Figure 24: Performance of TextBlob in classification of tickets

5 DISCUSSION OF RESULTS

In this chapter, the research findings presented in the previous chapters are discussed, in relationship to the research objectives and questions of this study. First, an explanation of the findings will be presented and then the accuracy and reliabilty of these finding will be investigated and presented. Conclusively, a presentation of deductions and hypotheses based on these research findings will be given as well as the challenges encountered during the study.

5.1 EXPLANATION OF RESEARCH RESULTS

CHANNELS OF CUSTOMER SUPPORT

It was discovered that WhatsApp was not only the social media platform on which most of the respondents have an account but also the one considered best through which one is to receive customer support. Many businesses use an email-first approach to providing customer support, which unfortunately may not be meeting the current consumer demands and requirements in this regard. In the same vein, respondents chose scheduled support and the ability to review and rate customer agents among the 4 most essential features in a trouble ticket system. Unfortunately, these are not provided in an external social media platform. Based on our quantitative findings too, it was seen that because social media platforms provide an unguided ticket creation process, they hinder the automation of various subsquent processes such as ticket triaging and routing as well as problem mining. This goes on to justify the necessity for a custom environment dedicated to managing the various processes involved in the pipeline of a trouble ticket in a trouble ticket system. Lastly, a significant number of tickets were found to contain private sensitive data such as customer phone numbers, emails, locations, and device details. This makes social media, a public platform, unsuitable for managing customer support in an ethical and 100% private manner, when necessary.

ESSENTIAL FEATURES AND FACTORS

According to our respomdents, a dedicated platform (website) was chosen as the most important feature in a trouble ticket system. This shows the high necessity for a platform that will meet the individual requirements of customers from a trouble ticket system. Consumers will need to know the state of the status of their trouble tickets besides getting recommendations of probable solutions. The least of attempts at implementing these features would necessitate a private environment built for and dedicated to customer support, both from the perspective of a customer and that of businesses and their customer support teams.

In addition to that, over $\frac{1}{3}$ of the analyzed tickets were found to require tier-0 support — support that does not require interaction with any support agent, but rather with self help material like a knowledge databank, a forum, articles and blog posts, among others. First, this shows the need for business and organizations to invest in self-help material. Also, it shows the need for a ticket classifier that can automatically direct such tickets to recommend solutions or self help material. Otherwise, if every ticket has a customer-agent interaction cost, the business ends up performing poorly across some metric of customer support quality.

CUSTOMER SUPPORT RELATIVE TO KPIs

The KPIs used to track customer support are very many. Our study primarily focused on the Average Reply Time (ART). A low positive correlation between the number of tickets handled by a client and their ART was found, which could imply that an opportunity exists, for the usage of AI to further reduce this correlation. It can be justified too that use an extra layer of channels to offer customer support reduces the customer effort score (CES). A manual investigation of the responses to the tickets further showed that the commonest phrase in the first reply was, *Send us the details in our inbox*, which goes on to show the inneficiency of using an eternal channel to collect customer tickets, and then direct them back to an internal channel.

PERFORMANCE BOOST BASED ON AI

PyABSA was seen to have a performance boost of over 4000%, assuming a customer support agent takes half a second to classify a ticket based on its aspects. This shows the ability of automation to improve the performance of a company's customer support offering across various KPIs such as Employee Satisfaction Score (ESAT), Ticket Volumes Serviced, Average Reolution Times (ART), Cost Per Resolution (CPR), among others. The model used was not effective enough in mining aspects, with a focus on comprehending jargon used in the telecom service sector. This was so because we used a pretrained model. Therefore, the successful implementation of such a model that understands telecom jargon will necessitate either prior collection of and training of the model on a similar dataset, or the usage and training of the model in-production.

5.2 CHALLENGES AND LIMITATIONS

The accuracy and effectiveness of solutions based on artificial intelligence are heavily reliant on the amount and form of data that is used to train the AI models. However, a lot of data from customer support interactions in businesses is usually reserved for internal use only. This makes it hard for third party entities to develop customized AI solutions for various commercial domains. In cases when the datasets are accessible, a lot of beuracracy and legal red tape has to be endured by the implementing party. In our case study, gaining access to the customer support datasets from both AIRTEL and MTN, despite being the preferred way, was expected to be time costly. Hence, we ended up using a dataset of only 1003 rows, out of a recommended 100,000 data rows, for our specific case study.

In addition to that, the models we used were pretrained on datasets that are not specifically annotated for jargon in the telecom industry. Hence, this accounts for some innacuracy in some performance results. Also, the pretrained model of the PyABSA library places a limitation to the size of text input (80 words) which could have accounted for some performance innacuracies too. The TextBlob library was found to have a somewhat low accuracy in performance too, which affected the quality of our findings, as far as sentiment analysis of the entire ticket (not aspects) was concerned.

5.3 HYPOTHESES BASED ON RESEARCH RESULTS

Twitter is and has been an effective channel for customer support. However, based on our findings, WhatsApp would be the best customer support channel, at least specifically for a client that matches our survey respondents' profiles. However, it was noted that the process of ticket creation can be either unguided or guided — input parameters are explicitly requested of th customer. Unfortunately, most social media channels offer an unguided ticket creation process. This leads to tickets having very low subjectivity. This affects the feasibility of automating the subsquent processses like triaging, routing, and problem mining using artificial intelligence. Because of that, a dedicated customer support platform would serve this purpose best, as it would be easier to customize the various processes, compared to social media channels that exist as external disparate systems.

6 CONCLUSION

Recent research in artificial intelligence has predominantly focused on the architecture and viability of various methods of implementation. However, in this study, we took a focus on the practical usability of a subdomain in artificial intelligence (natural language processing) in a day-to-day challenge in commerce in Uganda. Current research on the matter shows the practical effectiveness of using sentiment analysis to classify trouble tickets. However, this study has further expounded upon this ideaology, and proposed the usage of a dedicated platform, such as mobile application or website, to specifically collect and administer trouble tickets. In the same vein, a plethora of research has been conducted on this matter but with a greater focus on the international context. This study has provided a perspective to this from a Ugandan context.

Conclusively, based on the analysis conveyed, it can be concluded that Twitter is effective as a customer support channel but does it, together with other social media channels, do not present the most effective and efficient opportunities of automating various processes in th adminstration of customer support tickets. Furthermore, it has been justified that sentiment analysis has and can play a role in the automation of customer support systems. However, more detailed and accurate research studies will have to be conducted to fully justify this.

7 RECOMMENDATIONS

This section enlists the recommendations of the study. The purpose is to offer ideas on how the findings of this study can be implemented in both academia and commerce. We also offer suggestions on how the challenges discovered can be addressed for better outcomes.

The current study can be interpreted as a first step in the research on improving customer support in Uganda, specifically through leveraging artificial intelligence. However, the results of this study should also be treated with caution due to the rather small sample size of the respondents and trouble tickets used to justify the proposed method of automation through sentiment analysis. From a social perspective, more detailed studies should be undertaken to determine the perspectives on, as well as the social, legal and civil implications of implementanting them.

Further research could also examine the application of these proposed models in order to improve customer support in other subdomains, such as in e-governance and and ecommerce in general. Maintaining a major focus on the Ugandan context while conducting subsquent further research in this regard will immensely contribute to a deeper understanding of the state of customer support in Uganda. More anonymized customer support data should be made available by businesses, companies, and organizations to support private research and academic endavours at improving its adminstration.

8 REFERENCES

[Abola, 2016] Abola, J. (2016). Uganda: Culture of bad customer service. https://allafrica.com/stories/201610040005.html/.

[Barnecut, 2022] Barnecut, S. (2022). Ticket escalation: What it is and how to manage it. https://www.zendesk.com/blog/art-ticket-escalation-process/.

[Bennett et al., 2020] Bennett, V., Bowman, K., and Wright, S. (2020). Customer service benchmark report - telecommunication. Technical Report DOE-SLC-6903-1, Salt Lake City Corporation, Salt Lake City, UT.

[Bertram, 2009] Bertram, D. (2009). The social nature of issue tracking in software engineering. https://prism.ucalgary.ca/bitstream/handle/1880/104369/placeholder.pdf?sequence=1&isAllowed=n.

[Bhatti and Qureshi, 2007] Bhatti, K. K. and Qureshi, T. M. (2007). Impact of employee participation on job satisfaction, employee commitment and employee productivity. *International review of business research papers*, 3(2):54–68.

[BMC, 2022] BMC (2022). Cost per ticket: The ultimate service desk metric. https://www.bmc.com/blogs/cost-per-ticket/.

[Carlos, 2019] Carlos, W. (2019). Dissatisfied customer blasts dfcu bank over poor customer care. https://dailypostug.com/dissatisfied-customer-blasts-dfcu-bank-over-poor-customer-care/.

[Digivate, 2022] Digivate (2022). The pros and cons of using twitter for customer service - digivate.

[Farris et al., 2010] Farris, P. W., Bendle, N., Pfeifer, P. E., and Reibstein, D. (2010). *Marketing metrics: The definitive guide to measuring marketing performance*. Pearson Education.

[hiverHQ, 2022] hiverHQ (2022). Customer support and service – everything you need to know. http://hiverhq.com/customer-support-guide.

[Jindal and Liu, 2006] Jindal, N. and Liu, B. (2006). Mining comparative sentences and relations. In *Aaai*, volume 22, page 9.

[Kalyegira, 2021] Kalyegira, T. (2021). Poor customer care could be the country's greatest national crisis. https://www.monitor.co.ug/uganda/magazines/people-power/poor-customer-care-could-be-the-country-s-greatest-national-crisis-1514106.

[Kandel et al., 2011] Kandel, S., Heer, J., Plaisant, C., Kennedy, J., Van Ham, F., Riche, N. H., Weaver, C., Lee, B., Brodbeck, D., and Buono, P. (2011). Research directions in data wrangling: Visualizations and transformations for usable and credible data. *Information Visualization*, 10(4):271–288.

[Khder, 2021] Khder, M. A. (2021). Web scraping or web crawling: State of art, techniques, approaches and application. *International Journal of Advances in Soft Computing & Its Applications*, 13(3).

[Kravchuk and Schack, 1996] Kravchuk, R. S. and Schack, R. W. (1996). Designing effective performance-measurement systems under the government performance and results act of 1993. *Public administration review*, pages 348–358.

[Lakshmi and Mohideen, 2013] Lakshmi, S. and Mohideen, M. A. (2013). Issues in reliabilityand validity of research. *International journal of management research and reviews*, 3(4):2752.

[Leen, 2022] Leen, P. (2022). What is a ticketing system? — best 10 it ticketing software 2022. `https://www.itarian.com/ticketing-system.php`.

[Lin, 2022] Lin, Y. (2022). Top 10 customer service statistics for 2022 [new data] - oberlo. `https://www.oberlo.com/blog/customer-service-statistics`.

[Loria et al., 2018] Loria, S. et al. (2018). textblob documentation. *Release 0.15*, 2(8).

[Lucas, 2012] Lucas, R. W. (2012). *Customer service: Skills for success*. McGraw-Hill.

[Mancilla, 2022] Mancilla, E. (2022). The 5 levels of it help desk support. `https://blog.invgate.com/the-5-levels-of-it-support`.

[Mejova, 2009] Mejova, Y. (2009). Sentiment analysis: An overview. *University of Iowa, Computer Science Department*.

[Miles and Huberman, 1994] Miles, M. B. and Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. sage.

[Nord, 2022] Nord, T. (2022). What ticket routing means for your customer satisfaction. `https://www.ultimate.ai/blog/customer-experience/what-ticket-routing-means-for-your-customer-satisfaction`.

[Pang et al., 2008] Pang, B., Lee, L., et al. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in information retrieval*, 2(1–2):1–135.

[Paramesh et al., 2018] Paramesh, S., Ramya, C., and Shreedhara, K. (2018). Classifying the unstructured it service desk tickets using ensemble of classifiers. In *2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS)*, pages 221–227. IEEE.

[Radovic and Karapandzic, 2005] Radovic, M. and Karapandzic, S. (2005). Process engineering. *Faculty of Organizational Science, Beograd*.

[Sasaki and Kita, 1998] Sasaki, M. and Kita, K. (1998). Rule-based text categorization using hierarchical categories. In *SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 98CH36218)*, volume 3, pages 2827–2830. IEEE.

[Shimpi et al., 2014] Shimpi, V., Natu, M., Sadaphal, V., and Kulkarni, V. (2014). Problem identification by mining trouble tickets. In *Proceedings of the 20th International Conference on Management of Data*, pages 76–86.

[Singh, 2021] Singh, A. (2021). What are various service desk ticket types? `https://www.getapp.com/resources/what-are-various-service-desk-ticket-types/`.

[Statista, 2021] Statista (2021). Share of mobile subscriptions uganda 2015-2020, by operator. `https://www.statista.com/statistics/671666/mobile-subscription-share-in-uganda-by-operator/`.

[Tripp and Grégoire, 2011] Tripp, T. M. and Grégoire, Y. (2011). When unhappy customers strike back on the internet. *MIT Sloan Management Review*, 52(3):37–44.

[Twitter, 2022a] Twitter (2022a). Counting characters when composing tweets. https://developer.twitter.com/en/docs/counting-characters.

[Twitter, 2022b] Twitter (2022b). How twitter has become a key customer support channel.

[UCC, 2022] UCC (2022). Consumer rights. https://www.ucc.co.ug/consumer-rights/.

[Ward, 2019] Ward, C. (2019). 90% of consumers expect service queries to be resolved within 24hrs. https://www.mycustomer.com/service/channels/90-of-consumers-expect-service-queries-to-be-resolved-within-24hrs.

[Yang and Li, 2022] Yang, H. and Li, K. (2022). Pyabsa: Open framework for aspect-based sentiment analysis. *arXiv preprint arXiv:2208.01368*.